

## Individual on-line processing differences are not necessarily reflected in off-line acceptability judgments.

Dan Michel

University of California, San Diego

dmichel@ling.ucsd.edu

**Background:** Ross' (1967) introduction of island phenomena has triggered debates as to whether the source and nature of island phenomena are due to syntactic properties of the grammar (Sprouse, Wagers and Philips 2012a,b: henceforth SWP) or due more to general cognitive properties that underlie sentence processing (Hofmeister, Staum Casasanto & Sag 2012a,b: henceforth HSS). This recently resurgent debate is centered on the intuition, which we name the Cognitive Co-variation Intuition (CCI), that cognitive measures of individual differences should co-vary with offline acceptability judgments, here specifically applied to island phenomena as given in (1). We provide results from on-line and offline experiments that undermine this intuition.

- (1) If the unacceptability of island violations are due to processing difficulties arising from constraints on measurable cognitive resources (Kluender and Kutas 1993b) such as working memory (WM), then those with a measurably greater WM capacity are expected to process island violations more easily and thus rate them more acceptable.

However, SWP report no robust co-variation between WM measures and acceptability ratings of island violations, concluding that this processing account of islands was not supported. HSS voice various concerns with this conclusion, including whether (1) is a valid assumption.

**Experiments:** To address (1)'s validity, we conducted a self-paced reading (SPR) experiment (n=48) providing a measure of online processing difficulty and an acceptability judgment study (7-point scale, n=66) using identical materials (table 1). Participants completed cognitive measures of individual differences including a flanker task to assess selective attention, an n-back task to assess general working memory capacity and a verbal span task (Just and Carpenter, 1992) to assess verbal working memory. Finally, a memory-interference task assessed susceptibility to similarity-based interference. In this task subjects were given a series of words to remember and made a 'yes/no' response to a test list of words containing both words on the study list and words not on the study list, some of which were phonological and semantic lures.

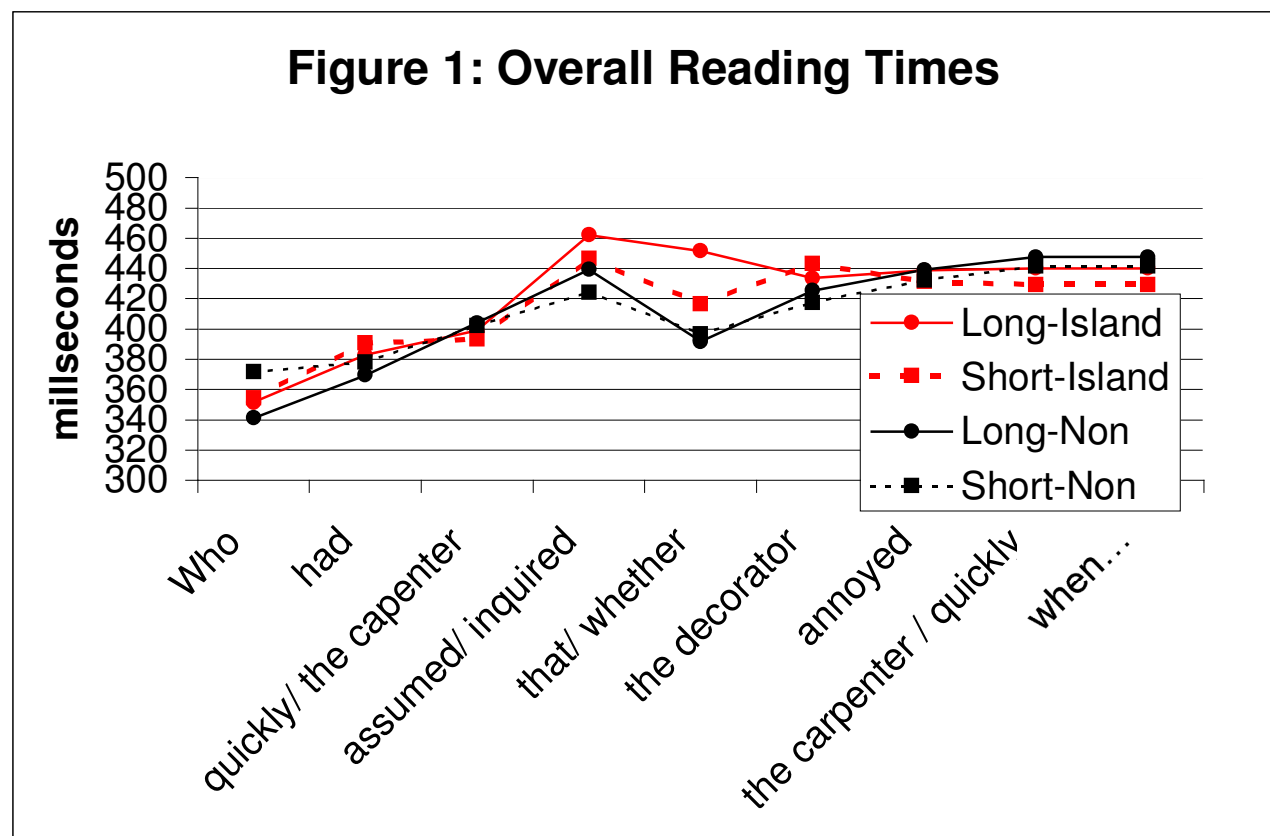
Following SWP (2012a), experimental items crossed structure: presence/absence of island structure, and length: whether extraction occurred out of the possible island (long) or not (short). A difference between the current materials and SWP (2012) is exemplified by the italicized word *quickly* and *the carpenter* in the example sentences in table 1. The insertion of the adverb at the gap location allowed for word position to be held constant throughout the remainder of the sentence. This was to avoid word position effects in the SPR experiment and an in progress ERP experiment. Additionally, these adverbs were frequency-matched with the noun of the NP that alternated with it in the gap/NP sentence position. This allowed for a frequency-controlled comparison in the self-paced reading experiment. A norming study was done which showed that the addition of these adverbs to the sentences did not alter the acceptability ratings given to them. Additionally, 168 filler items of various complexity, acceptability and length were added to the 32 experimental items to form four lists using a Latin-square design, additionally balanced for order. High and low-scoring groups for each cognitive measure were formed by median split. Groups, conditions and word positions (for SPR) were submitted to ANOVAs.

Table 1: Example sentences. **Mean acceptability** (out of seven); **RT** (measured on bold word)

	Non-island	Island
Short	(A)Who had _ <i>quickly</i> assumed <b>that</b> the decorator annoyed <i>the carpenter</i> when the deadline was missed? <b>acceptability</b> : 4.70; <b>RT</b> : 396.9ms	(B)Who had _ <i>quickly</i> inquired <b>whether</b> the decorator annoyed <i>the carpenter</i> when the deadline was missed? <b>acceptability</b> : 4.54; <b>RT</b> : 416.7ms
Long	(C)Who had <i>the carpenter</i> assumed <b>that</b> the decorator annoyed _ <i>quickly</i> when the deadline was missed? <b>acceptability</b> : 3.33; <b>RT</b> : 391.8ms	(D)Who had <i>the carpenter</i> inquired <b>whether</b> the decorator annoyed _ <i>quickly</i> when the deadline was missed? <b>acceptability</b> : 2.79; <b>RT</b> : 451.6ms

**Results and Discussion:** In the acceptability judgment task subjects rated short extraction sentences higher than long extraction sentences, non-island sentences were rated higher than island sentences, and the island violation condition (D) was rated lowest of any of the conditions. (Values in Table 1)

The overall results of the self-paced reading experiment show a three-way distinction at the clause boundary (*that/whether*) with the island violation condition (D: extraction out of an island structure) read the slowest, followed by the short extraction-island condition (B). Finally, there was no statistical difference between the two conditions read most quickly at the clause boundary (both the short extraction-non-island (A) and long extraction-island conditions (C)). No other significant differences were found in the sentence, including at the gap site. (See table 1 and figure 1)



Turning to the individual differences effects, we find that the high scorers on the recall-interference task (those who are *less* susceptible to similarity-based interference) show a pattern not predicted by the CCI (1). High scorers rate the difference between the long and short extraction sentences as significantly greater than low scorers. Additionally, high scorers *slow down* at the clause boundary compared to low scorers. (All  $p < 0.05$ , values in table 2)

Table 2: Recall-interference and extraction distance:

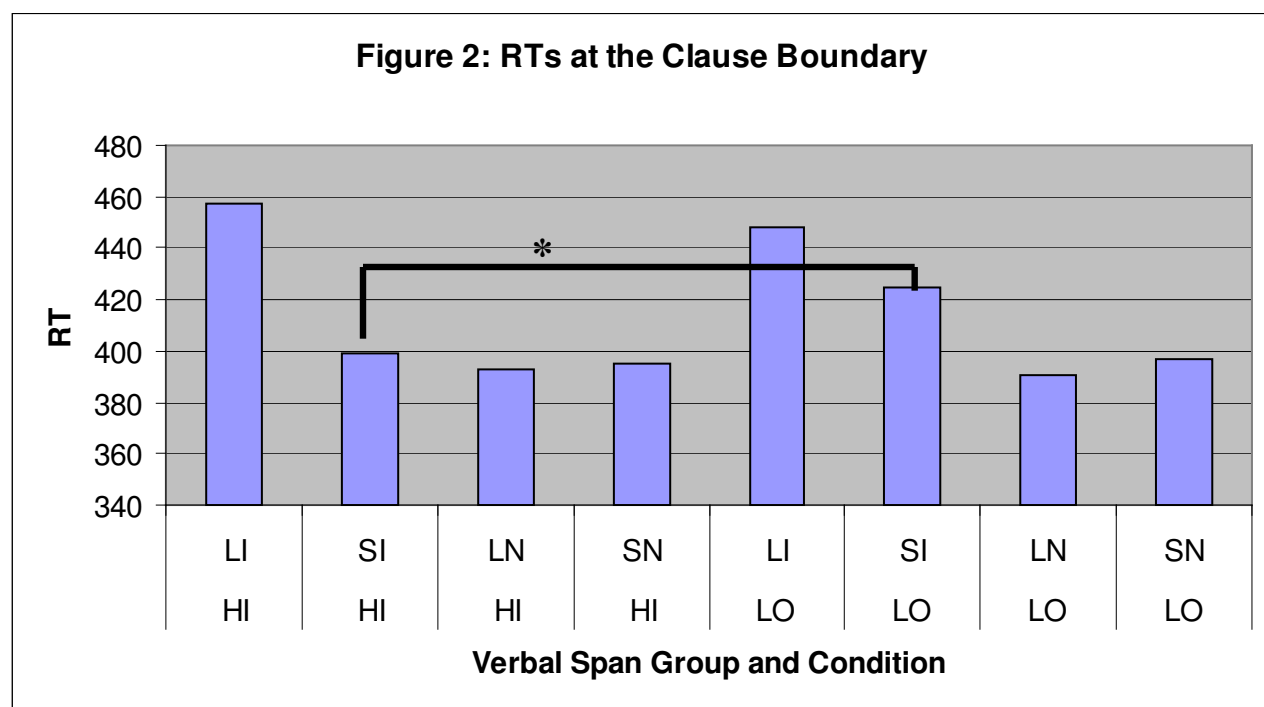
**Mean acceptability** (out of seven); **RT** (at clause boundary *that/whether*)

	High-scoring	Low-scoring
Short	<b>acceptability:</b> 5.10 <b>RT:</b> 427.0ms	<b>acceptability:</b> 4.23 <b>RT:</b> 390.9ms
Long	<b>acceptability:</b> 3.12 <b>RT:</b> 447.9ms	<b>acceptability:</b> 3.06 <b>RT:</b> 398.5ms

While slower reading times are often interpreted as a sign of processing difficulty, we argue that in this case they reflect an additional effort being made by the high scorers that does not necessarily imply difficulty. This interpretation is available due to the pattern found in the acceptability judgment data. It appears that the high scorers are slowing down at the clause boundary to more carefully distinguish between the long and short extraction conditions. This pattern of careful reading is not predicted by the CCI.

Results from the high and low scorers on the verbal span task also do not support the CCI. First, like SWP (2012), we find no significant evidence that individuals with higher WM capacity rate the island violations higher than individuals with lower WM capacity. This is true whether the verbal span or n-back task is used, whether raw or z-scores are used, and whether the size of the interaction (akin to SWP's DD score) or the ratings of the island-violating sentence (D in table 1) are taken in isolation. This replicates SWP's findings and expands upon them with additional WM measures and analyses. In the self-paced reading, there is a pervasive main effect of the higher scoring individuals reading faster overall, but an interaction with conditions is found only at the clause boundary position.

In Figure 2 we see the RTs broken down by condition and group. 'Hi' and 'lo' indicate high and low scorers on the verbal span task, respectively. Above this, 'I' indicates the island condition and 'N' indicates non-island. 'S' and 'L' likewise indicate short and long extraction such that LI is the long-island violation condition. While it is unsurprisingly read the slowest, there is no significant difference between the high and low groups in this condition. Similarly there is no statistical difference detected in the non-island conditions between or within the WM groups. The only processing advantage found for the high span group is in the short-island condition. In the high WM group, this condition is not significantly different from the non-island conditions. That is, the high WM group does not show a cost just for the presence of an island structure. This is not true for the low WM group where RTs slow from the non-island conditions to the short island and again from the short island to the long island (violation) condition (all  $p < 0.05$ ). As the high group still has a(n equal) processing cost to the island violation, we can infer that the advantage to the high group is limited to being able to process the island structure more easily only if there is no island violating dependency remaining.



This finding is contra the CCI (1) for two reasons. First, it indicates the lack of a relationship between the online processing of islands and their acceptability scores with respect to grouping individuals by WM scores. Second, it reveals that high WM individuals do not process island violations themselves any better than low WM individuals, but only the clause boundary.

The findings here indicate that the CCI is untenable if applied to islands on a macro level. However, the fact that co-variation of individual cognitive measures with the building blocks of island violations, namely extraction and structure, opens the possibility that the CCI can be adjusted and applied on a less broad level of analysis.

## References:

- Hofmeister, P., Staum Casasanto, L., & Sag, I. A. (2012a) How do individual cognitive differences relate to acceptability judgments? A reply to Sprouse, Wagers, and Phillips. *Language* 88, 390-400.
- Hofmeister, P., Staum Casasanto, L., & Sag, I. A. (2012b) Misapplying working-memory tests: A reductio ad absurdum. *Language* 88, 408-409.
- Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review* 98, 122–149.
- Kluender, R., & Kutas, M. (1993a). Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5 (2), 196-214.
- Ross, J. R. (1967). Constraints on Variables in Syntax. PhD Thesis. MIT.
- Sprouse, J., Wagers, M., & Phillips, C. (2012a). A test of the relation between working memory capacity and syntactic island effects. *Language* 88, 82-123.
- Sprouse, J., Wagers, M., & Phillips, C. (2012b). Working-memory capacity and island effects: A reminder of the issues and the facts. *Language* 88, 401-407.